☑ | Generate Collection | | Print |

L3: Entry 1 of 29   File: PGPB   Feb 13, 2003

DOCUMENT-IDENTIFIER: US 20030033477 A1
TITLE: Method for raid striped I/O request generation using a shared scatter gather list

Summary of Invention Paragraph:
[0007] When the host operating system sends an I/O request to move data between the disk and its memory, it provides a scatter gather list (SGL) that describes the actual memory regions to move the data to or from. Typically, the memory regions are fragmented and data is scattered to the various regions when data is transferred to host memory space. When memory is retrieved from host memory space and written to a drive, it is gathered from the various memory regions to form scatter gather list. Typically, striping is implemented by taking the Scatter Gather List (SGL) for the transfer data, which provides a description of the location of the data, and formulating a new SGL for the data transfer to each drive. The original operating system host I/O request is divided into several new I/O requests to be sent to each of the stripe disk members and break up the original scatter gather list is divided into several smaller lists (one for each new I/O request). This method requires extra processing and memory resources to parse the original SGL and create new smaller ones dedicated for each of the new I/O requests.

Detail Description Paragraph:
[0024] The gathered data is assembled as a scatter gather list, as per step 120. One of the types of SGL, the chain SGL, may be useful where the data is located in noncontiguous physical memory. In general, each SGL entry contains an address and a length and may contain flags, such as Size of Address (i.e., 32-bit or 64-bit), End of List Reached, direction of data transfer, and the like. The SGL may be a set of pointers containing address locations which point to the stored bytes and information concerning the stored bytes, a data structure containing an initial address, address offsets, and length or start and stop addresses for each of the initial address and address offsets, and the like.

Detail Description Paragraph:
[0026] In step 140, the single scatter gather list remains intact during processing. Data for each physical I/O request is extracted as needed. The data extraction per physical I/O request may be performed through hardware, software, and/or firmware. The single scatter gather list may reside in host memory or may be copied to disk array controller memory.

Detail Description Paragraph:
[0030] FIGS. 3A and 3B illustrate an SGL list. In FIG. 3A, an SGL entry 220 may contain a flag field. The flag field bits may be dedicated for certain parameter information, such as end of list, address size, direction read/write, local/system address, entry or element type (i.e., simple or chain), end of buffer, and last element. FIG. 3B shows an example of a scatter gather list 230 having five simple element SGL entries 220 that describe a total of 256K bytes.

Detail Description Paragraph:
[0032] FIG. 4 illustrates another embodiment of the present invention. In this embodiment, the data may be very fragmented in host memory. The resulting scatter

gather list may be quite large to fully account for the data to be transferred
through a host I/O request to accommodate the multiple start addresses and either
byte lengths or end addresses which correspond to each data stripe. In the method,
the SGL entry number n and stripe number m are initialized, per step 500. Data is
transferred according to the SGL entry number n and the stripe number m, per step
505. A determination is made if all the data has been transferred which is
referenced by SGL entry n, step 510. If it has not been fully transferred, it is
transferred to stripe m, per step 515. Otherwise, a determination is made as to
whether the end of the SGL has been reached, per step 530. If the end of the SGL
has not been reached and the data defined by the last entry of the SGL has been
fully transferred, the next entry in the SGL is accessed, per step 540. Otherwise,
transfer for the single host I/O request is ended, per step 535. If data is
transferred as defined by SGL entry n, as per step 515, a determination is then
made as to whether the end of stripe m has been reached, per step 520. If it has,
then another determination is made whether this is the last stripe, step 550. If it
is, then transfer for the host I/O request ends, per step 555. Otherwise, the next
stripe m is accessed, per step 545. The data is then written to or read from stripe
m, per step 525, and processing returns to step 510. Variations of this method are
contemplated by the present invention. For example, step 515 may follow step 520 or
steps 515 and 525 may be merged into a single transfer step. Data transfer may
occur directly between the host device and the disk arrays or it may be temporarily
stored within the disk array controller memory. Further, an initial determination
may be made after step 500 as to whether there is actually data defined by the SGL,
whether the disk drives are full, or whether there is enough space in a data stripe
for the data. Also, step 505 may be merged into step 515 and/or step 525.